

**GEOGRAPHY OF CRETACEOUS EXTINCTIONS: DATA BASE
DEVELOPMENT**

D. M. Raup*
University of Chicago

Data bases built from the source literature are plagued by problems of data quality. Unless the data acquisition is done by experts, working slowly, the data base may contain so much "garbage" that true signals and patterns cannot be detected. On the other hand, high quality data bases develop so slowly that satisfactory statistical analysis may never be possible due to the small sample sizes. This report describes results of a test of the opposite strategy: rapid data acquisition by non-experts with minimal control on data quality.

186 published lists of species and genera of fossil invertebrates of latest Cretaceous age (Maestrichtian) were located through a random search of the paleobiological and geological literature. The geographic location for each faunal list was then transformed electronically to Maestrichtian latitude and longitude and the lists were further digested to identify the genera occurring in each ten-degree, latitude-longitude block. The geographical lists were clustered using the Otsuka similarity coefficient and a standard unweight-pair-group method. The resulting clusters are remarkably consistent geographically, indicating that a strong biogeographic signal is visible despite low-quality data.

A further test evaluated the geographic pattern of end-Cretaceous extinctions. All genera in the data base were compared with Sepkoski's compendium of time ranges of genera to determine which of the reported genera survived the Cretaceous mass extinction. In turn, extinction rates for the ten-degree, latitude-longitude blocks were mapped. The resulting distribution is readily interpretable as a robust pattern of the geography of the mass extinction.

The study demonstrates that a low-quality data base, built rapidly, can provide a basis for meaningful analysis of past biotic events.